AFRL-RI-RS-TR-2009-104
Final Technical Report
April 2009

REPRESENTATION AND ANALYSIS OF PROBABILISTIC INTELLIGENCE DATA (RAPID)

Carnegie Mellon University

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th ABW, Wright-Patterson AFB Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RI-RS-TR-2009-104 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/ /s/

NANCY A. ROBERTS
Work Unit Manager

JOSEPH CAMERA, Chief
Information & Intelligence Exploitation Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
APR 09	Final	Jun 07 – Jan 09

4. TITLE AND SUBTITLE

REPRESENTATION AND ANALYSIS OF PROBABILISTIC INTELLIGENCE DATA (RAPID)

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA8750-07-2-0137

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)

Eugene Fink, Jaime G. Carbonell, Anatole Gershman,
Ganesh Mani, Dwight Dietrich

5d. PROJECT NUMBER
PAIN

5e. TASK NUMBER
00

5f. WORK UNIT NUMBER
10

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Carnegie Mellon University
5000 Forbes Ave
Pittsburgh PA 15213-3815

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/RIED
525 Brooks Rd.
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFRL-RI-RS-TR-2009-104

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. PA# 88 ABW-2009-1390

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Tools were developed for the representation and analysis of uncertainty in INTEL data and targeted uncertainty reduction. The purpose is to help INTEL analysts answer these questions: 1) What hypotheses can be validated/refuted based on available uncertain data and at what level of certainty? 2) What missing data is critical for verifying or refuting given hypotheses and increasing the certainty of current conclusions? 3) What are the tradeoffs between the value of specific missing data and cost and difficulty of obtaining it?

15. SUBJECT TERMS
Uncertainty, probability, reasoning with uncertainty, proactive intelligence, prediction, design strategies, systems architecture, leadership modeling, test probes

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Nancy A. Roberts
U	U	U	UU	27	19b. TELEPHONE NUMBER (*Include area code*) N/A

Table of Contents

Executive Summary 1
Background and Problems Addressed 2
Informal Example 3
Objectives and Performance Goals 5
Technical Approach 7
Experimental Evaluation 10
Accomplishments 15
Recommendations for Future Research 17
References 19
List of Acronyms 22

List of Tables

Table 1: Experimintal results when treating variables as if independent 11
Table 2: Best expected values..... 12
Table 3: Prior is set to 70%..... 13
Table 4: Results of selection of the best model 13
Table 5: Results when trying to distinguish 1st model from rest. 14

Executive Summary

Analysts must cope with uncertain and partial information, for such is the nature of intelligence analysis; however, little support is provided by present-day software for reasoning with uncertain numerical values (e.g. quantities, dates, and locations), uncertain nominal values (e.g. identities of individuals), or uncertain causative or associative relations (e.g., who reports to whom and how strictly). We have investigated the related problems and developed a prototype system with the following capabilities:

- Representation of uncertainty as probability density functions.
- Reasoning with uncertainty in order to determine relative likelihood of outcomes.
- Proactive prioritization of further intelligence gathering for the purpose of reducing uncertainty and reliably discriminating among given hypotheses.

For instance, if assessing the nuclear capabilities of a hostile country, there may be many uncertain observations: the cumulative nuclear-physics expertise, the quantity of obtained fissionable material, the location of nuclear research facilities, the identity of the key players, external networks of clandestine procurement, and the allocated budget. Reducing one uncertainty (e.g. ascertaining the quantity and type of fissionable materials) may prove more important than another (e.g. the exact number of scientists on the project) in determining whether the country can quickly develop nuclear weapons, and therefore worth greater efforts (e.g. risking a human intelligence asset). The developed prototype tools helps to make such determinations.

Functionally, we have built software tools that reason with uncertainties as probability density functions over observables, with the objectives of determining whether there is enough information to discriminate among competing hypotheses; if not, exactly what new information should be gathered.

The specific role of these tools within the Proactive Intelligence (PAINT) architecture includes the quantitative evaluation of the likelihood of given hypotheses, as well as the information value of given observations and probes. Our tools provide this information to the *Probe Strategy* module, which is another component of PAINT, delivered by Lockheed Martin, and the *Probe Strategy* module uses it to construct intelligence collection plans. Our main software products and related deliverables for the PAINT program are as follows:

- Developed a library of data structures and procedures for probability computations based on uncertain data, represented by probability density functions.
- Developed probabilistic mechanisms for evaluating the likelihood of given hypotheses and information value of observations and probes, and adapted them to the needs of the PAINT architecture.
- Developed an Excel-based graphical user interface for the stand-alone use of our tools, which allows the viewing and editing of uncertain data.
- Packaged the developed tools for the integration into the PAINT architecture, and for their stand-alone use through Excel.

Background and Problems Addressed

When analysts process intelligence data, they usually have to work with uncertain and incomplete information. Standard software tools enforce the use of exact values or ranges, and provide almost no support for describing uncertainty, such as a probability distribution conditioned on thus-far available data. Nor does analytic software, to the best of our knowledge, provide for automated re-computing of said probability distributions upon receipt of new data. The long-term purpose of our work is to address these limitations, develop a general-purpose system for the processing of uncertain data, and integrate it with standard data-management tools, such as Excel and Oracle.

Before the beginning of the PAINT program, we investigated techniques for representing uncertain nominals and numeric values as part of the Reflective Agents with Distributed Adaptive Reasoning project (RADAR) under Defense Advanced Research Projects Agency, and reported the results in a series of papers [Fink *et al*., 2006a; Fink *et al*., 2006b; Bardak *et al*., 2006a; Bardak *et al*., 2006b]. The developed tools allowed a human administrator to describe uncertain data related to volatile crisis situations, analyze its implications, and construct plans for coping with the evolving crisis. We also investigated techniques for analyzing massive intelligence data during work on the Audit Record Generation and Utilization System project under Disruptive Technologies Office/Advanced Research and Development Activity, and developed mechanisms for the indexing of massive structured data and fast retrieval of data that provided approximate matches to given queries [Fink *et al*., 2004a; Gazen *et al*., 2004; Carbonell *et al*., 2005; Carbonell *et al.*, 2006; Jin and Carbonell, 2006]. On the negative side, the ARGUS tools did not support explicit representation of uncertainty, which limited their applicability in the analysis of uncertain and volatile situations.

During the PAINT work, we have continued this research and investigated the problem of developing a general-purpose system for the representation and analysis of structured uncertain data; this system has been named RAPID. The system represents and updates uncertainty in a principled manner, permits inference over uncertain data, and suggests what new intelligence would be maximally definitive in uncertainty reduction. The latter capability is the most novel capability of RAPID; knowing what information would be useful to gather provides a proactive aspect to analysis.

Informal Example

We begin with a simple example that requires distinguishing between two hypotheses. We consider the task of an economic analyst who is observing a small pharmaceutical company and trying to infer the plans of its management.

The company has recently advertised the sale of a new medication, and it claims that its main focus is on expanding its production. The analyst however suspects that the company is working on another, more advanced medication, which has a potential for greater sales. If she is right, the company works on this new medication in secrecy, which is understandable, since it does not want to reveal its plans to the competitors.

The analyst is trying to decide whether she is right based on available public data, along with some private data gained through her "special" channels, such as talking with her friends in pharmaceutical industry and hearing private opinions of fellow analysts. Thus, she had to distinguish between two mutually exclusive models:

- M_1: The company focuses on the production of its current medication.
- M_2: The company puts significant resources into development of a new medication.

The analyst also has to account for the possibility that neither of her hypotheses is correct, and something entirely different may be in the offing. For instance, the company may switch to production of medical equipment or it may file of a bankruptcy. While the chances of these unexpected outcomes are low, they are not negligible.

We suppose further that the analyst has some idea of prior probabilities from her past experience with such situations. For instance, she may believe that the prior probability of M_1 is 0.6, that of M_2 is 0.3, and the chances that neither hypothesis is correct are 0.1.

If she had no other data, she would use these probabilities; however, she has other data, which include public accounting numbers provided by the company, statements of its president, announced contracts with other companies, news about recent pharmaceutical developments, and so on. While a lot of these data may be irrelevant or very inaccurate, some may turn out to be "gold nuggets" that would greatly help in her task. The analyst has to identify relevant data, evaluate the chances of each model, and decide which additional data she should gather to improve her evaluation. While these tasks may sound similar to standard Bayesian reasoning, there are several important differences, which lead to a novel challenging problem:

- Available data may be correlated in complex ways; the analyst may not know about these correlations, and she cannot reliably determine which data are dependent on each other. For instance, two news articles may come from different sources (making them near-independent) or from the same source (making them highly dependent), and the analyst does not know which is the case.
- The analyst may have a lot of data, for instance if the company is in the news and under public scrutiny. Or she may have very little data, for instance, if it is a new small company. In either case, she must make the best use of the available data, and make as accurate conclusions as she can.

- The analyst has to plan collection of additional data, which cannot be addressed by standard Bayesian techniques.

Objectives and Performance Goals

We next describe the inputs and outputs of the RAPID module in the PAINT architecture, and the metrics for measuring its performance. We have *not* yet evaluated its performance according to these metrics since we have only recently completed building the system. We are now continuing this investigation beyond the end of the PAINT program, and we expect to have quantitative measurements in the next several months.

Inputs and outputs

The RAPID module of the PAINT architecture receives the following inputs:

- Probabilistic pathway and leadership models, provided by the modeling components of the PAINT architecture.
- Available external observations of the target scenario, which may include measures of the observation uncertainty.
- A list of hypotheses that need to be confirmed, refuted, or sharpened.
- A list of possible future observations and related active probes, as well as the prior probability distributions of these observations under each hypothesis, if known.

RAPID evaluates the likelihood of each given hypothesis, ranks the hypotheses by their likelihood, and outputs the ranked list along with the probability of each hypothesis. It also evaluates the information value of potential future observations and probes. This information value is based on the expected reduction of the information entropy and related notion of the Kullback Leibler (KL)-distance. The system evaluates the expected entropy reduction for each possible observation and each probe, and selects the observations and probes with the greatest expected reduction. It ranks the observations and probes by their value, and outputs their ranked list along with the value estimates. It provides this information to the *Probe Strategy* module of PAINT, which then uses the resulting estimates to construct information-gathering strategies.

Performance metrics

The problem of building general-purpose systems for reasoning under uncertainty has not yet received much attention from the computer science community, and researchers have not developed standard performance metrics for such systems. While researchers have built a number of special-case systems (for example, see a review of such systems in a recent manuscript by Bardak, Fink, and Carbonell [Bardak *et al*., 2009]), they have not addressed a general case. The review of the previous work has revealed that the existing special-case metrics are not applicable to our PAINT work, and we have developed two new general-purpose metrics, both on the scale from −1.0 (worst) to 1.0 (best). The first shows the accuracy of evaluating the likelihood of given hypotheses, and the second is for the accuracy of evaluating the information value of observations and probes. We also outline metrics for evaluating the system scalability.

Likelihood of given hypotheses: We measure the accuracy of hypothesis evaluation by comparing RAPID's ranking of hypotheses with the ground-truth ranking. Specifically, suppose that RAPID needs to rank n different hypotheses, and their ground-truth ranking from the most likely to the least likely is H_1, H_2,…, H_n. Suppose further that RAPID

outputs a ranking $G_1, G_2,\ldots,G_n$, which includes the same hypotheses in a different order. The number *inv* of inversions required to convert the output order $G_1, G_2,\ldots, G_n$ into the correct order $H_1, H_2,\ldots, H_n$ serves as the raw measure of the output quality. This value corresponds to the number of transpositions in the Bubble-Sort algorithm. The best possible value of this raw measure is 0, which would mean that the output order is perfectly correct. On the other hand, the expected value for a *random* ordering, which does not account for any available data, is $n\,(n-1)\,/\,4$. To normalize the raw measure to the [−1.0, 1.0] scale, we use the following expression as the final metric:

$$\textbf{Equation 1:}\quad 1-\frac{inv}{n(n-1)/4}.$$

For a random ordering, the expected value of this normalized metric is 0.0; for the worst possible ranking, it is −1.0; and for the perfectly correct ranking, it is 1.0. In practice, we never expect to under-perform the random ordering, so the effective range is [0.0, 1.0]. *The goal has been to achieve the initial performance of about 0.5 according to this metric, and then extend the system to reach the 0.9–1.0 range.*

Information value of observations and probes: The metric for evaluating potential observations and probes is analogous to that for hypothesis evaluation. We compare RAPID's ranking of observations and probes by their information value with the ground-truth ranking, and use Equation 1 as the final accuracy metric. *The goal has been to achieve about the same performance as for the hypothesis evaluation; that is, to build an initial system with performance of about 0.5, and then extend it to reach the 0.9–1.0 range.*

Scalability: We use two metrics for evaluating the scalability of RAPID. The first is the maximal number of hypotheses that can be evaluated within a one-minute time limit, and the second is the maximal number of potential observations and probes that can be ranked within the same one-minute limit. *The goal has been to achieve the initial performance of about 100, and then build a far more scalable system, which will process 100,000 hypotheses and probes in a minute.*

Note that this scalability metric is only for evaluating the probability computations within our module, and it does not include the time required for running the related simulations by modeling components. The evaluation of the modeling time is a separate problem, which is not directly related to the performance requirements of our module. In the current version of the PAINT architecture, the computational time of our module is about two orders of magnitude smaller than the modeling time. Thus, it is not on a critical path for improving the overall performance of the PAINT architecture.

Technical Approach

Uncertain data: We have developed a mechanism for encoding uncertain data, which supports the representation of uncertain nominal values, strings, numbers, math functions, and dependencies among the available data. We represent an uncertain nominal or string by a probability distribution over possible values, and an uncertain integer or real number by a distribution over possible ranges. Furthermore, we allow specification of uncertain mathematical functions by piecewise-linear functions with uncertain coordinates of segment endpoints, and by probability distributions over multiple possible functions.

We have implemented a prototype library that supports the application of all standard arithmetic and logical operations to uncertain values. For instance, it allows computing the sum of uncertain values, applying an uncertain function to an uncertain value, and determining the probability that two uncertain strings are identical. It also allows qualitative representation of uncertainty, such as "completely unknown" values.

We have also developed a basic language for describing dependencies among uncertain data by inference rules, based on the extension of the inference-rule mechanisms developed during the work on PRODIGY [Veloso *et al.*, 1995; Fink and Blythe, 2005] and RADAR [Fink *et al.*, 2006a; Fink *et al.*, 2006b]. We represent each inference rule by its preconditions and effects. The system supports the use of all standard arithmetic and logical operations, as well as user-specified mathematical functions, in defining rule effects.

Identification of critical uncertainties: The system determines which of the uncertain or missing data has the greatest impact on the completion of current tasks, such as discriminating among hypotheses, and which additional data would help to complete these tasks. For each task, the system identifies the relevant data, evaluates the impact of each related uncertainty on the task completion, and ranks these uncertainties by their impact. The evaluation procedure uses game-theoretic search through possible data-collection scenarios. That is, it considers potential data-collection strategies and all possible outcomes of each strategy, determines the impact and likelihood of each outcome, and identifies the strategy that is most likely to lead to the task completion.

We do not make any assumptions about the volume of available information, and the described technique works both in data-rich and data-poor situations. If we have a large volume of data, the technique helps identify the most important indicators among the available information. If we have only a few observations, the technique makes the best use of this limited data.

We now give a more detailed description of the uncertainty-analysis procedures integrated into the PAINT architecture. We assume that the target of information gathering is some organization, such as a business, terrorist group, or hostile government. We are interested in a specific aspect of the organization, called *intent*, which can take one of several mutually exclusive values $M = \{M_1 \ldots M_n\}$. These values are our hypotheses, and we have a prior probability distribution Q over them. An example of such hypotheses may be the intent of an organization to enter certain markets, which is not directly observable. We can observe several other aspects of the organization $\{X_i\}$, each of which can take several discrete values. For example, we may observe the purchases of certain machinery and hiring of specialists. We denote the set of all observations by O.

The organization has several *control points* that we can influence through probes. For example, we may be able to influence the ability of the organization to obtain funding or hire certain specialists. Each control point is represented by a variable whose value affects the organization behavior. We have a black-box model of the organization, which takes the control point settings $\{R_i\}$ and the hypotheses as parameters, and produces the probability distributions of the observable variables. Thus, we are provided with conditional probability distributions $P(X_i \mid M_j, R_k)$.

The developed RAPID tools perform three functions:

- Quantitative assessment of the information value of control-point settings.
- Quantitative assessment of the information value of knowing more precisely the value of an individual control-point variable.
- Revision of the hypothesis probability estimates based on new observations.

To describe how we are accomplishing the first two tasks, we first define the notion of information value. This value of an observation is not an absolute quantity; it depends on what we previously knew and what we may do differently as a result of the new observation. Suppose that we have a set of available actions Φ. Each action $\alpha \in \Phi$ incurs different costs $C_\alpha(M_i)$ under different conditions $M_1, \ldots, M_n$. For example, our action may be to buy some commodity in an anticipation of the target organization's entry into a specific market. If it does, our final costs will be low; if not, the costs will be high. Given a probability distribution Q over M, we can anticipate the expected cost of this action:

Equation 2: $\hat{C}_\alpha(Q) = \sum_k C_\alpha(M_k) \cdot P(M_k)$

Given Q, a rational decision maker would select the course of action $\Phi(Q)$ that minimizes $\hat{C}_{\Phi(Q)}(Q)$. Instead of costs, we can use utilities, which are the negation of the costs.

We can define the value of information that leads us to believe that M has a different distribution Q' instead of Q. If we do not get this information, we select actions based on Q, but they incur costs according to Q'. These costs are no smaller than the costs incurred by the optimal actions selected based on Q'. The difference is the information value:

Equation 3: $V(Q',Q) = \hat{C}_{\Phi(Q)}(Q') - \hat{C}_{\Phi(Q')}(Q')$

Thus, in order to determine the value of information that changes our beliefs from Q to Q', we need a decision model that defines the expected costs of acting based on Q in the world described by Q'.

To compare it to the more traditional definition of information gain, consider the following model. We have two hypotheses, M_1 and M_2, which stand for hostile vs. benign intentions of our adversary. Our decision involves allocating a proportion $t \in [0, 1]$ of our budget to military preparations. The cost of our decision is $-\log t$ in case of M_1 and $-\log (1-t)$ in case of M_2. Thus, if we spend no money and the adversary turns out hostile, the cost is infinitely high. The same is true if we spend our entire budget and the adversary turns out benign. Assuming that under Q, $P(M_1) = p$ and under Q', $P(M_2) = p'$, the expected cost of the decision under Q is minimized when $t = p$ and under Q' when $t = p'$. Then, the expected cost $\hat{C}_{\Phi(Q')}(Q')$ of the optimal decision $\Phi(Q')$ under Q' is

Equation 4: $\hat{C}_{\Phi(Q')}(Q') = -p' \log p' - (1-p') \log (1-p') = H(Q')$,

where $H(Q')$ is Shannon's entropy. The expected cost $\hat{C}_{\Phi(Q)}(Q')$ of the decision $\Phi(Q)$ under Q' is

Equation 5: $\hat{C}_{\Phi(Q)}(Q') = -p' \log p - (1-p') \log (1-p) = H(Q', Q)$,

where $H(Q', Q)$ is Shannon's cross entropy and the information value $V(Q', Q)$ is

Equation 6: $V(Q', Q) = H(Q', Q) - H(Q') = D_{KL}(Q' \| Q)$,

where $D_{KL}(Q' \| Q)$ is the KL divergence between the prior and the posterior probability distribution, which is also known as information gain.

In real applications, the costs of our decisions may not be as drastic as in the above example, and the information value function may not be as mathematically convenient as the KL divergence. Our definition of information value is a generalization of the traditional information gain. We assume that the domain experts supply us with a decision model that defines cost functions. If we do not have this expert advice, we use the standard information gain.

We now show how to compute the expected information value of a control point. The the PAINT modeling components produce conditional probability distributions $P(X_i \mid M_j, R_k)$ for each observation variable X_i. First, we discuss the case of a single observation variable $O = \{o_i\}$. For each value o_i, we compute the posterior probability distribution $Q'(o_i, R_k)$ of M and obtain the information value of o_i:

Equation 7: $U(o_i, R_k) = V(Q'(o_i, R_k), Q)$

$Q'(o_i, R_k)$ is computed by calculating $P(M_j \mid o_i, R_k)$ for all j.

Equation 8: $P(M_j \mid o_i, R_k) = P(o_i \mid M_j, R_k) \cdot P(M_j, R_k) / P(o_j, R_k)$

Since the intent M and the control point setting R are independent, we conclude that

Equation 9: $P(M_j \mid o_i, R_k) = P(o_i \mid M_j, R_k) \cdot P(M_j) / \sum_n P(o_j \mid M_n, R_k) \cdot P(M_n)$

All quantities in the above formula are known: $P(o_i \mid M_j, R_k)$ come from the PAINT model and $P(M_j)$ are the prior beliefs. Thus, given the information value function V, we can compute $V(o_i)$ for the control point R_k. The expected information value $\breve{U}$ for R_k is

Equation 10: $\breve{U}(R_k) = \sum_i [U(o_i, R_k) \cdot P(o_i \mid R_k)] = \sum_i [U(o_i, R_k) \cdot \sum_n [P(o_j \mid M_n, R_k) \cdot P(M_n)]]$

This approach works only when we have a single observation. In reality, we may have thousands of observation variables. Combining them into one and obtaining its joint probability distribution is impractical. Therefore, we use an approximate solution. For each control point setting R_k, we select the observation variable X_i that maximizes the expected information value of R_k. Then $\breve{U}_i(R_k)$, which is the expected information value based on observing only X_i, becomes our lower-bound approximation of $\breve{U}(R_k)$. If we make an actual observation under R_k conditions, we use the value of X_i to compute the resulting posterior distribution of M.

Experimental Evaluation

The experiments have two objectives: (1) the assessment of methods for calculating information value of data potentially obtained through intelligence gathering and (2) the assessment of methods for calculating the information value of probes.

We have experimented with four PAINT models: a benign model M1 and three nefarious models M2, M3, and M4. Model M2 contains all segments of Model M1; Model M3 contains all segments of Model M2; and Model M4 contains all the segments of Model M3. As a result, M2 is the closest (most difficult to distinguish) to M1, while M4 is the furthest. M1 contains 43 segments, M2 contains 50 segments, M3 contains 53 segments, and M4 contains 57 segments. We assume that intelligence reports provide the percent completion information for two different weeks for each of the 42 benign segments, giving us 84 observation variables. We also assume that the values of these variables are rounded to fit into 22 bins: bin 1 if the segment has not been started; bin 2 if the percent completion is between 0 and 5%; bin 3 between 5% and 10%, and so on; bin 22 corresponds to the completed segment. We have eliminated all nefarious segments with no influence on the observation variables, which explains the relatively small number of the remaining segments.

Each segment specifies the expected progress for each week, which is proportional to the allocated resources. We treat the expected progress as the mean and multiply it by a normally distributed "jitter" with the standard deviation 3% (we also tried other values of the standard deviation), which introduces considerable non-determinism. We consider four probes, that is, modifications of resource allocations, which have different effects on each model. The use of probes increases the total number of different models to 20.

Each simulation run of the model produces 84 bin values. We have run 11,000 simulations to get estimates of the bin probabilities for each model. We have computed 999 trailing averages of 10,000 simulations. The standard deviation of the averages is less than .002 for all variables.

Given an intelligence report, the system assesses the likelihood of the underlying model. Since the dimensionality of the problem (84 variables) makes the direct calculation of the posterior probabilities intractable, we need to use an approximate method. We have tested ten such methods by running 2000 simulations. In each simulation, we have used each model to produce an intelligence report containing 84 bin values. Based on this report, each algorithm selects the most likely underlying model. The score for each algorithm is defined as the proportion of incorrect model selections.

To test the volatility of the data, we have run 3000 simulations and computed 999 trailing score averages of 2000 simulations. The standard deviation of the averages for the most common information-gain algorithm is less than .005, which means that score differences in excess of 3% are statistically significant. The intelligence-report data points with probability less than the "credulity threshold" of .02 are excluded. If all data points of a report are below the credulity threshold, it is a good indication of the "none of the above" case, which means that the system rejects all given models.

The first set of experiments involves pairwise comparisons of models. We tested the following selection algorithms:

- **Ind:** all variable are treated as if they were independent

- **Rand:** we select one representative variable at random and compute the posteriors as if it were the only variable
- **CG** ("center of gravity"): for each model, we compute the posteriors for each variable and take the average
- **Exp IG:** before looking at an intelligence report, we select the variable that gives the highest expected information gain; we then use its value from the report
- **Exp Ent:** as above, we select the variable that gives us the lowest expected entropy
- **Exp IV:** as above, we select the variable that gives us the highest expected information value based on the cost function of the selection decision
- **Exp JSD:** as above, we select the variable that gives us the highest expected Jensen-Shannon Divergence between the distributions of binary probabilities of the two models
- **IG:** we compute the posteriors for each variable in the report and then select the one with the highest information gain (KL divergence) between the posterior and the prior
- **Ent:** as above, we compute the posteriors for each variable and then select the one with the lowest entropy
- **IV:** as above, we compute the posteriors for each variable and then select the one with the highest information value based on the cost function of the selection decision

The following table shows the experimental results. The notation m1_3, for example, indicates Model 1 with modifications caused by Probe 3. The prior probability of the benign model is 50%.

Table 1: Experimintal results when treating variables as if independent

Prior=.5	indep	rand	cg	exp ig	exp ent	exp iv	exp jsd	ig	ent	iv
m1_1 vs m2_1	0.50	0.47	0.48	0.40	0.40	0.40	0.40	0.40	0.40	0.40
m1_2 vs m2_2	0.51	0.43	0.15	0.01	0.01	0.01	0.01	0.00	0.00	0.00
m1_3 vs m2_3	0.50	0.49	0.51	0.21	0.21	0.21	0.21	0.21	0.21	0.21
m1_4 vs m2_4	0.49	0.46	0.33	0.11	0.11	0.11	0.11	0.13	0.13	0.13
m1_5 vs m2_5	0.49	0.47	0.38	0.20	0.20	0.20	0.20	0.18	0.18	0.18
m1_1 vs m3_1	0.51	0.46	0.36	0.16	0.16	0.16	0.16	0.17	0.17	0.17
m1_2 vs m3_2	0.50	0.40	0.13	0.01	0.01	0.01	0.01	0.00	0.00	0.00
m1_3 vs m3_3	0.50	0.48	0.47	0.21	0.21	0.21	0.21	0.19	0.19	0.19
m1_4 vs m3_4	0.50	0.46	0.30	0.11	0.11	0.11	0.11	0.13	0.13	0.13
m1_5 vs m3_5	0.51	0.45	0.30	0.16	0.16	0.15	0.16	0.10	0.10	0.10
m1_1 vs m4_1	0.50	0.41	0.18	0.00	0.00	0.00	0.00	0.00	0.00	0.00
m1_2 vs m4_2	0.51	0.40	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00
m1_3 vs m4_3	0.50	0.41	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00
m1_4 vs m4_4	0.51	0.41	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00
m1_5 vs m4_5	0.51	0.39	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The table shows that treating the variables as if they were independent is not a good idea; in this case, the algorithm cannot distinguish between any two models. Random selection

of the representative variable does a little better because sometimes a good variable is selected by chance. The “center of gravity” algorithm is better yet, but not nearly as good as the rest of the algorithms.

The next four algorithms select a representative variable prior to obtaining intelligence reports. As expected, when the prior probability of the benign model is 50%, all four algorithms give essentially the same answer. More surprisingly, no significant differences appear when the prior is 70%, although the performance of all algorithms deteriorates by 1–10% when the mix between the two models remains fifty-fifty.

Because these four algorithms select a representative variable prior to obtaining the intelligence reports, they can be used to rank the available probes. The table below shows the best expected values of the measures computed by the four algorithms. The ranking of the probes is exactly the same.

Table 2: Best expected values

Prior=.5	**variable**	**exp ig**	**exp ent**	**exp iv**	**exp jsd**
m1_1 vs m2_1	60	0.04	0.04	0.10	0.04
m1_2 vs m2_2	18	0.96	0.96	0.49	0.96
m1_3 vs m2_3	38	0.35	0.35	0.27	0.35
m1_4 vs m2_4	70	0.75	0.75	0.39	0.75
m1_5 vs m2_5	46	0.45	0.45	0.29	0.45

The last three algorithms select the “best” posteriors of all variables based on the calculation of information gain, entropy, and information value. Again, as expected, when the prior probability of the benign model is 50%, all three algorithms give essentially the same answers. More surprisingly, the answers are essentially the same as from the four representative variable algorithms. When the prior is set to 70%, some differences between the three algorithms begin to appear without systematically favoring any one of them as shown in the table below.

Table 3: Prior is set to 70%

Prior=.7	exp ig	exp ent	exp iv	exp jsd	ig	ent	iv
m1_1 vs m2_1	0.49	0.50	0.50	0.50	0.49	0.50	0.50
m1_2 vs m2_2	0.01	0.01	0.01	0.01	0.01	0.00	0.19
m1_3 vs m2_3	0.30	0.30	0.30	0.30	0.30	0.43	0.30
m1_4 vs m2_4	0.14	0.14	0.14	0.14	0.10	0.14	0.11
m1_5 vs m2_5	0.34	0.33	0.34	0.34	0.19	0.33	0.20
m1_1 vs m3_1	0.23	0.23	0.23	0.23	0.16	0.25	0.16
m1_2 vs m3_2	0.01	0.01	0.01	0.01	0.04	0.00	0.20
m1_3 vs m3_3	0.29	0.29	0.29	0.29	0.23	0.41	0.25
m1_4 vs m3_4	0.13	0.13	0.13	0.13	0.11	0.10	0.13
m1_5 vs m3_5	0.21	0.21	0.21	0.21	0.11	0.12	0.13
m1_1 vs m4_1	0.00	0.01	0.00	0.00	0.00	0.00	0.25
m1_2 vs m4_2	0.00	0.01	0.00	0.00	0.08	0.00	0.33
m1_3 vs m4_3	0.01	0.01	0.00	0.00	0.00	0.00	0.23
m1_4 vs m4_4	0.00	0.01	0.00	0.00	0.00	0.00	0.25
m1_5 vs m4_5	0.00	0.01	0.00	0.00	0.00	0.00	0.28

The second set of experiments addresses the selection of the best model among four candidates. The following table shows the results.

Table 4: Results of selection of the best model

Prior=.5	exp ig 4	exp iv 4	ig 4	iv 4
Probe 1	0.51	0.50	0.39	0.57
Probe 2	0.51	0.39	0.24	0.46
Probe 3	0.50	0.53	0.45	0.57
Probe 4	0.50	0.51	0.34	0.55
Probe 5	0.50	0.50	0.31	0.52

Since the available data are insufficient for a reliable selection, the algorithms make a number of wrong conclusions; however, the results are much better than random. The information-gain algorithm performs somewhat better than the rest.

The final set of experiments also involves four models, but we are now trying to distinguish between the first model and the rest, which is an easier task. The following table shows the results.

Table 5: Results when trying to distinguish 1st model from rest.

Prior=.5	**exp ig 1+3**	**exp iv 1+3**	**ig 1+3**	**iv 1+3**
Probe 1	0.24	0.23	0.22	0.25
Probe 2	0.06	0.06	0.12	0.19
Probe 3	0.14	0.14	0.14	0.17
Probe 4	0.06	0.06	0.08	0.13
Probe 5	0.11	0.11	0.13	0.16

Accomplishments

Technical accomplishments: The role of the RAPID tools within PAINT is to evaluate the likelihood of a given hypotheses and the information value of observations and probes. In the diagram below, we show the overall PAINT architecture and the position of our tools within this architecture, which is the *RAPID Hypothesis Classifier* component at the bottom. The main technical accomplishments of our work have been as follows:

- Developed a mechanism for the representation of uncertain structured data by probability density functions and if-then inference rules; it supports nominal values, numeric values, strings, and dependencies.
- Developed a library of procedures for probability computations based on uncertain data.
- Developed probabilistic mechanisms for evaluating the likelihood of hypotheses and information value of observations and probes, and adapted them to the needs of the integrated PAINT architecture.
- Developed an Excel-based graphical user interface for the stand-alone version of the RAPID tools, which allows the viewing and editing of uncertain data.
- Packaged the RAPID tools for the integration into the PAINT architecture and for the stand-alone use through Excel.

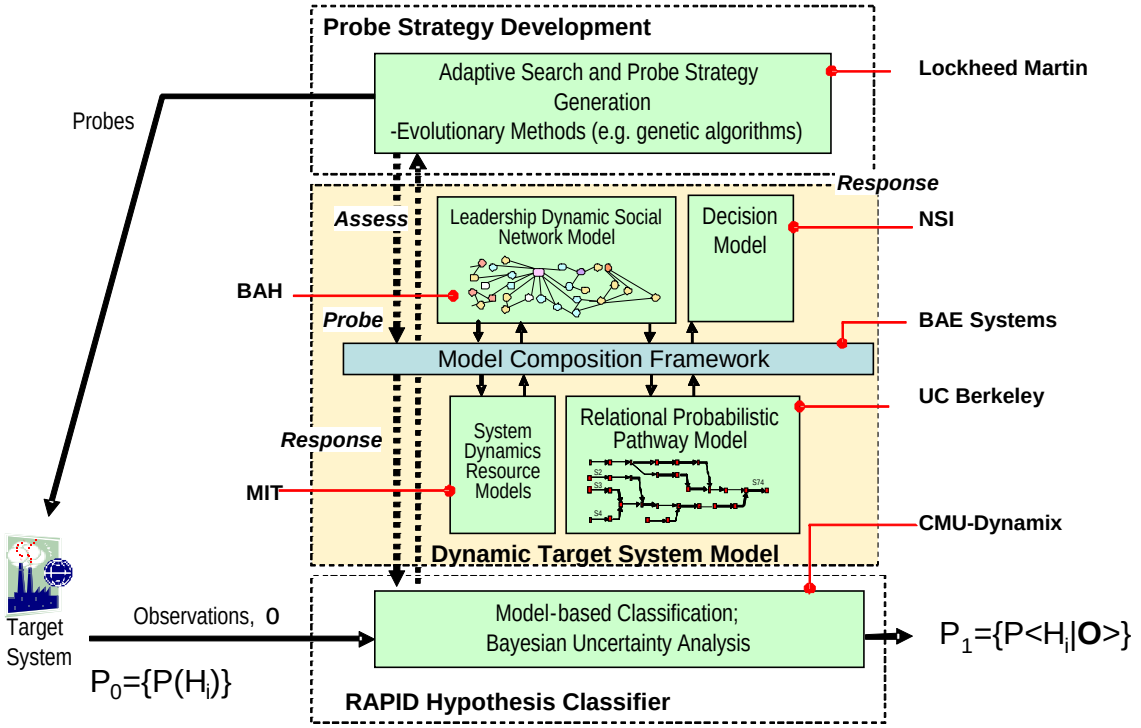

Figure 1: PAINT Architecture

Products: We have delivered two main products for the integration into the PAINT architecture, which are part of the *RAPID Hypothesis Classifier* box in the diagram:

- *Hypothesis evaluation:* This module is responsible for evaluating the likelihood of each given hypothesis; it accounts for the prior probabilities and new uncertain evidence. It also evaluates the likelihood that none of the given hypotheses are correct, thus detecting "surprise" situations.
- *Evaluation of observations and probes:* This module is responsible for evaluating the information value of specific observations and active probes. The evaluation results serve as input to the *Probe Strategy* module, which is the top box on the diagram.

We have integrated these two modules with the PAINT architecture. We have also developed a stand-alone version integrated with Excel, which allows the use of the standard Excel functionality with the RAPID uncertainty analysis tools.

Publications: We have published a conference paper based on the initial work under the PAINT program:

- Bin Fu, Eugene Fink, and Jaime G. Carbonell. Analysis of uncertain data: Tools for representation and processing. In Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, 2008.

We are currently working on two more conference papers, which will describe the developed strategies for the hypothesis evaluation and probe selection.

Recommendations for Future Research

We believe that our work under PAINT can be continued in two main directions. The first is the construction of general-purpose tools for the analysis and integration of uncertain data, and the second is the application of these tools to specific practical tasks.

General-purpose tools for the analysis and integration of uncertainty

A major long-term challenge is to develop a suite of general-purpose software tools for the viewing, analysis, and integration of uncertain data, which will help military and business analysts automate routine tasks involved in the evaluation of uncertain factors. The concept behind these tools is similar to the concept of a spreadsheet for the analysis of complex numeric data. The tools will enable analysts to build task-specific applications for the uncertainty analysis, quickly propagate basic reasoning results through large-scale datasets, and obtain visual representation of these results, much in the same way as Excel helps build and view task-specific numeric applications. In other words, the challenge is to build an advanced Excel-like spreadsheet package for the uncertainty processing, what-if analysis, and planning of additional data collection.

This challenge includes the development of mechanisms for the representation and fast processing of structured uncertain data, and integration of these mechanisms into the standard Excel software. The resulting system should enable military analysts, who may not have programming experience or advanced math background, to process complex uncertain data and build related spreadsheet applications for new tasks, in the same way as users without programming experience build task-specific Excel sheets. In particular, it should support the following capabilities:

- Representation of incomplete and uncertain data, including error margins, min-max ranges, sets of possible values, probability distributions, and qualitative uncertainty.
- Analysis of given hypotheses, evaluation of the certainty of specific conclusions, and identification of critical missing data related to specific reasoning steps.
- Identification of important contingencies, semi-automated construction of related what-if scenarios, and evaluation of their impact on the overall reasoning.
- Semi-automated integration of data from various sources, which may have different levels of accuracy, reliability, and “softness”; for example, integrated use of sensor data, human intelligence reports, and expert opinions.
- Continuous real-time update of the situation assessment based on a (possibly massive) stream of newly incoming data.

The related work is likely to involve the following research and engineering challenges, although this list may not be complete:

- Develop fast scalable algorithms for all main operations on uncertain data, and ensure that the speed of uncertain-data processing and related inference propagation is close to the normal speed of spreadsheet computations.
- Build a suite of tools for the identification of critical uncertainties, semi-automated interactive analysis of if-then scenarios, and planning of additional data collection.

- Design a general-purpose Application Programming Interface (API) for the integration of the proposed system with standard database systems, such as Oracle, and with streams of incoming new data; ensure the scalability of processing massive data streams through this API.
- Implement Graphical User Interface tools that enable analysts without programming experience to use the proposed system.

Specialized applications for the analysis of uncertain data

Although researchers have investigated a number of situation-awareness models, such as social networks, production pathways, and management decision processes, they have done almost no work on the explicit modeling of related uncertainties. In other words, researchers usually develop fully deterministic models, based on the implicit assumption that they can ignore the related uncertainties without a major loss of the reasoning accuracy. This traditional assumption may be a legacy of standard software applications, such as Excel and Oracle, which provide little support for the analysis of uncertainty and what-if scenarios.

On the other hand, human experts who deal with practical problems have accumulated much evidence on the importance of explicit reasoning about uncertainty and contingencies. For example, military analysts often point out that even the best-laid plans tend to go awry when they do not account for the current and future uncertainties.

The work under the PAINT program has confirmed that deterministic models are often insufficient for an accurate evaluation and prediction of complex systems, and that analysts need advanced tools for the explicit uncertainty analysis. Specifically, our experiments with production pathway and resource models, which were used in evaluating the uncertainty-analysis tools, have shown that even simple models exhibit a "chaotic" behavior in most cases, which means that small input changes often cause drastic output changes. Thus, if we do not account for uncertainty and use deterministic models with deterministic inputs, then we often get uninformative results, since minor changes to the models or their inputs, which are likely in uncertain situations, would invalidate most conclusions. The PAINT modeling teams have reported similar observations on their more advanced pathway, resource, and social-network models.

To address this need, we have developed a suite of prototype data structures and algorithms for the analysis of uncertain data and evaluation of its impact on the certainty of specific conclusions. A related future work direction is to integrate the developed algorithms with several special-purpose applications and demonstrate that they improve the effectiveness of these applications. This effort is likely to involve the following three research and engineering challenges.

- Extend the initial data structures and algorithms for the processing of uncertain data; improve their scalability and enhance the related contingency analysis.
- Develop practical applications for the uncertainty processing and what-if analysis, by integrating these algorithms with earlier deterministic-analysis applications.
- Integrate the resulting applications with mechanisms for the processing of massive streams of uncertain data.

Many of these interesting challenges will be addressed in future developments of proactive intelligence and reasoning under uncertainty.

References

- [Agrawal *et al*., 1993] Rakesh Agrawal, Tomasz Imielinski, and Arun N. Swami. Mining association rules between sets of items in large databases. In Proceedings of the ACM International Conference on Management of Data, pages 207–216, 1993.
- [Ballard, 1983] Bruce W. Ballard. The *-minimax search procedure for trees containing chance nodes. Artificial intelligence, 21(3), pages 327–350, 1983.
- [Bardak *et al*., 2006a] Ulas Bardak, Eugene Fink, and Jaime G. Carbonell. Scheduling with uncertain resources: Representation and utility function. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 1486–1492, 2006.
- [Bardak *et al*., 2006b] Ulas Bardak, Eugene Fink, Chris R. Martens, and Jaime G. Carbonell. Scheduling with uncertain resources: Elicitation of additional data. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 1493–1498, 2006.
- [Bardak *et al*., 2009] Ulas Bardak, Eugene Fink, and Jaime G. Carbonell. Scheduling with uncertain resources: Information elicitation. Unpublished manuscript, 2009.
- [Carbonell *et al*., 2005] Jaime G. Carbonell, Eugene Fink, Chun Jin, B. Cenk Gazen, Santosh Ananthraman, Philip J. Hayes, Ganesh Mani, and Dwight Dietrich. Exploring massive structured data in ARGUS. In Proceedings of the NIMD Principal Investigator Meeting, 2005.
- [Carbonell *et al.*, 2006] Jaime G. Carbonell, Eugene Fink, Chun Jin, B. Cenk Gazen, Johny Mathew, Abhay Saxena, Vini Satish, Santosh Ananthraman, Dwight Dietrich, and Ganesh Mani. Scalable data exploration and novelty detection. In Proceedings of the NIMD Principal Investigator Meeting, 2006.
- [Fikes and Nilsson, 1993] Richard E. Fikes and Nils J. Nilsson. STRIPS, a retrospective. Artificial Intelligence, 59(1–2), pages 227–232, 1993.
- [Fink *et al*., 2004a] Eugene Fink, Aaron Goldstein, Philip J. Hayes, and Jaime G. Carbonell. Search for approximate matches in large databases. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 1431–1435, 2004.
- [Fink *et al*., 2004b] Eugene Fink, Josh Johnson, and Jenny Hu. Exchange market for complex commodities: Theory and experiments. Netnomics, 6(1), pages 21–42, 2004.
- [Fink *et al*., 2006a] Eugene Fink, Ulas Bardak, Brandon Rothrock, and Jaime G. Carbonell. Scheduling with uncertain resources: Collaboration with the user. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 11–17, 2006.
- [Fink *et al*., 2006b] Eugene Fink, P. Matthew Jennings, Ulas Bardak, Jean Oh, Stephen F. Smith, and Jaime G. Carbonell. Scheduling with uncertain resources: Search for a near-optimal solution. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pages 137–144, 2006.
- [Fink *et al*., 2006c] Eugene Fink, Jianli Gong, and Josh Johnson. Exchange market for complex commodities: Search for optimal matches. Journal of Experimental and Theoretical Artificial Intelligence, to appear.

- [Fink and Blythe, 2005] Eugene Fink and Jim Blythe. Prodigy bidirectional planning. Journal of Experimental and Theoretical Artificial Intelligence, 17(3), pages 161–200, 2005.
- [Gazen *et al.*, 2004] B. Cenk Gazen, Jaime G. Carbonell, Philip J. Hayes, Chun Jin, and Eugene Fink. Hypothesis formation and tracking in ARGUS. In Proceedings of the NIMD Principal Investigator Meeting, 2004.
- [Ghahramani, 1998] Zoubin Ghahramani. Learning dynamic Bayesian networks. In C. Lee Giles and Marco Gori, editors, Adaptive Processing of Sequences and Data Structures, pages 168–197. Springer-Verlag, Berlin, Germany, 1998.
- [Guttman, 1984] Antonin Guttman. R-trees: A dynamic index structure for spatial searching. In Proceedings of the SIGMOD Conference, pages 47–57, 1984.
- [Hart *et al.*, 1968] Peter E. Hart, Nils J. Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics, SSC-4(2), pages 100–107, 1968.
- [Jin and Carbonell, 2006] Chun Jin and Jaime G. Carbonell. ARGUS: Efficient scalable continuous query optimization for large-volume data streams. In Proceedings of the Tenth International Database Engineering and Applications Symposium, 2006.
- [Jordan *et al*, 1994] Michael I. Jordan and Robert A. Jacobs. Hierarchical mixtures of experts and the EM algorithm. Neural Computation, 6, pages 181–214, 1994.
- [Jordan *et al*., 1997] Michael I. Jordan, Zoubin Ghahramani, and Lawrence K. Saul. Hidden Markov decision trees. In Michael C. Mozer, Michael I. Jordan, and Thomas Petsche, editors, Advances in Neural Information Processing Systems 9. MIT Press, Cambridge, MA, 1997.
- [Jordan *et al.*, 1999] Michael I. Jordan, Zoubin Ghahramani, Tommi S. Jaakkola, and Lawrence K. Saul. An introduction to variational methods for graphical models. Machine Learning, 37(2), pages 183–233, 1999.
- [Korf, 1985] Richard E. Korf. Depth-first iterative-deepening: An optimal admissible tree search. Artificial Intelligence, 27(1), pages 97–109, 1985.
- [Korf, 1995] Richard E. Korf. Space-efficient search algorithms. ACM Computing Surveys, 27(3), pages 337–339, 1995.
- [Liu *et al*., 2005] Yan Liu, Jaime G. Carbonell, Peter Weigele, and Vanathi Gopalakrishnan. Segmentation conditional random fields (SCRFs): A new approach for protein fold recognition. In Proceedings of the Ninth Annual International Conference on Research in Computational Molecular Biology, pages 408–422, 2005.
- [Liu *et al*., 2007] Yan Liu, Jaime G. Carbonell, and Vanathi Gopalakrishnan. Protein quaternary fold recognition using conditional graphical models. In Proceedings of the Twentieth International Joint Conference on Artificial Intelligence, 2007.
- [MacKay, 1999] David J. C. MacKay. Introduction to Monte Carlo methods. In Michael I. Jordan, editor, Learning in Graphical Models. MIT Press, Cambridge, MA, pages 175–204, 1999.
- [Mitchell, 1997] Tom M. Mitchell. Machine Learning. McGraw-Hill, New York, NY, 1997.
- [Morrison, 1968] Donald R. Morrison. PATRICIA—practical algorithm to retrieve information coded in alphanumeric. Journal of the ACM, 15(4), pages 514–534, 1968.

- [Orenstein, 1982] Jack A. Orenstein. Multidimensional tries used for associative searching. Information Processing Letters, 14(4), pages 150–157, 1982.
- [Parisi, 1998] Giorgio Parisi. Statistical Field Theory. Perseus Books Group, Jacksonville, TN, 1998.
- [Parunak and Brueckner, 2006] H. Van Dyke Parunak and Sven A. Brueckner. Extrapolation of the opponent's past behaviors. In A. Kott and W. McEneany, editors, Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind. CRC Press, Boca Raton, FL, pages 49–76, 2006.
- [Rabiner, 1989] Lawrence R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of IEEE, 77(2), pages 257–285, 1989.
- [Russell, 1992] Stuart J. Russell. Efficient memory-bounded search methods. In Proceedings of the Tenth European Conference on Artificial Intelligence, pages 1–5, 1992.
- [Russell and Norvig, 2003] Stuart J. Russell and Peter Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River, NJ, second edition, 2003.
- [Stonebraker *et al.*, 1986] Michael Stonebraker, Timos K. Sellis, and Eric N. Hanson. An analysis of rule indexing implementations in data base systems. In Proceedings of the First International Conference on Expert Database Systems, pages 465–476, 1986.
- [Tesauro, 1995] Gerald Tesauro. Temporal difference learning and TD-Gammon. Communications of the association for computing machinery, 38(3), pages 58–68, 1995.
- [Veloso *et al*., 1995] Manuela M. Veloso, Jaime G. Carbonell, M. Alicia Perez, Daniel Borrajo, Eugene Fink, and Jim Blythe. Integrated planning and learning: The PRODIGY architecture. Journal of Experimental and Theoretical Artificial Intelligence, 7(1), pages 81–120, 1995.
- [Viterbi, 1967] Andrew J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Transactions on Information Theory, 13(2), pages 260–269, 1967.
- [Weiner, 1973] Peter Weiner. Linear pattern matching algorithm. In Proceedings of the Fourteenth IEEE Annual Symposium on Switching and Automata Theory, pages 1–11, 1973.

List of Acronyms

API – Application Programming Interface
KL - Kullback Leibler
PAINT – ProActive INTelligence
RADAR - Reflective Agents with Distributed Adaptive Reasoning project
RAPID – Representation and Analysis of Probabilistic Intelligence Data